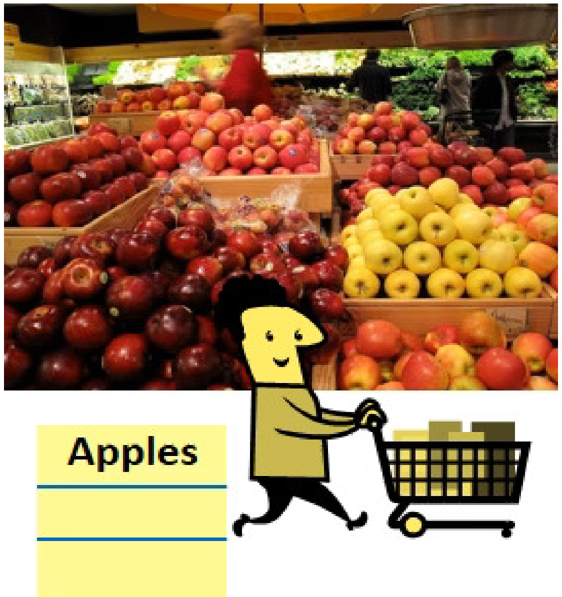



# Enabling Physical Analytics in Retail Stores Using Smart Glasses

**Swati Rallapalli**[*]

*Joint work with*

**Aishwarya Ganesan**[+], **Krishna Chintalapudi**[+], **Venkat Padmanabhan**[+], **Lili Qiu**[*]

[*]*The University of Texas at Austin* [+]*Microsoft Research India*

**ACM MobiCom 2014**

**September 8, 2014**

Page Views
Clicks
Purchases

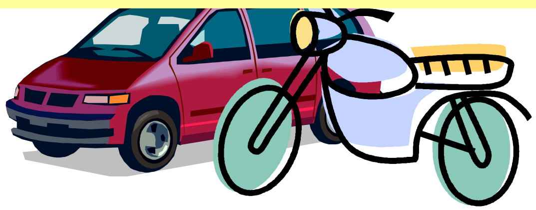
Web Analytics

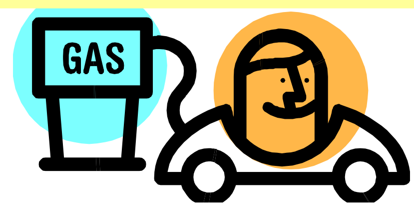
Location
Demographics
Interests
...

Targeted Ads

3x growth in 5 years from 2009 - 2013

**Important to capture shopper behavior not only in the online world but also in the physical world**





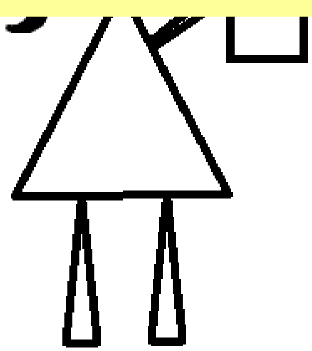

Even excluding those categories:

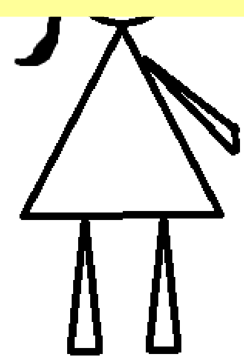
Online sales in U.S. in 2013 only

11%


**Human behavior in the physical world:**

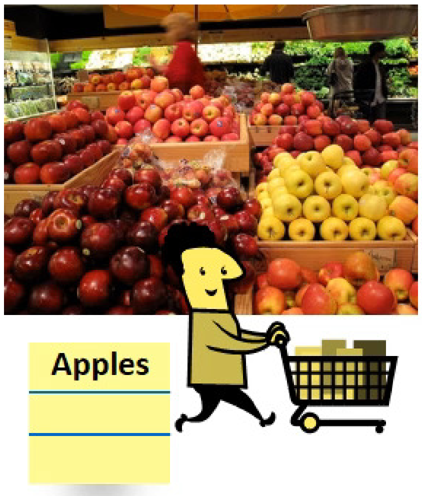
- ✓ **Contains signals missed by web-analytics & loyalty programs**
- ✓ **Indicates reinforced interest in buying**





## Physical Analytics

**Understanding the intent of the shoppers in the physical world**




## Benefits


**Alice sees a coupon as she is about to walk away!**

- **Enable contextual recommendations**
- **Shopping list reminders**
- **Guides to product locations**

# Our approach

Phytics Engine



**Enable a wide coverage and obtain rich user profiles!**


**Do not tie ourselves closely to any store..**

Technology

Localization, Product layouts, User analytics

Incentives

**Stores:** increased sales

**Physical analytics provider:** share of profits by partnering with stores

**Users:** discounts, shopping

Privacy


Survey: Co

Mar 14, 2014 at

Share 37

A new consum
provider Swirl,
and Ani and Tin
indoor location
consumer acce
overblown.

It is true that co
to their location
clear value for
doing so.

January 17, 2012, 9:00am EST

Shoppers Willing to Tell All

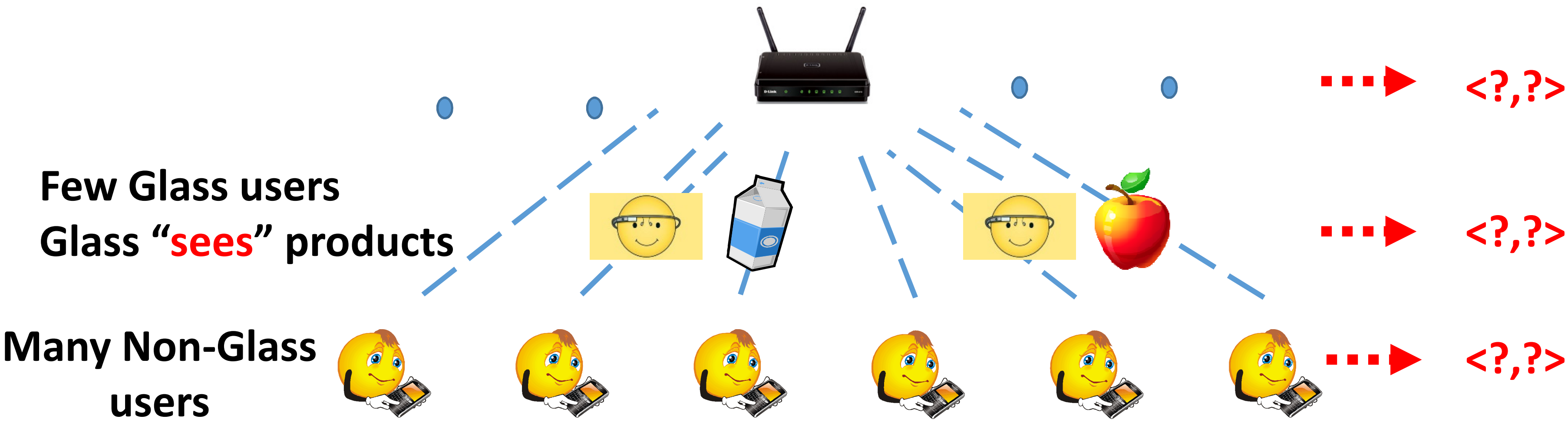
Teresa Novellino
Upstart Business Journal Entrepreneurs & Enterprises Editor
Email | Twitter

It might surprise retailers, but a new IBM study reveals that consumers are much more willing to give up information about themselves.

AutoLayout

Behavior Classification

# Overview: ThirdEye - AutoLayout




**Key Idea:**

**Wi-Fi**

+

**Inertial sensor**

+

**Video**

**Offline: AutoLayout**

**Localize Glass users**

+

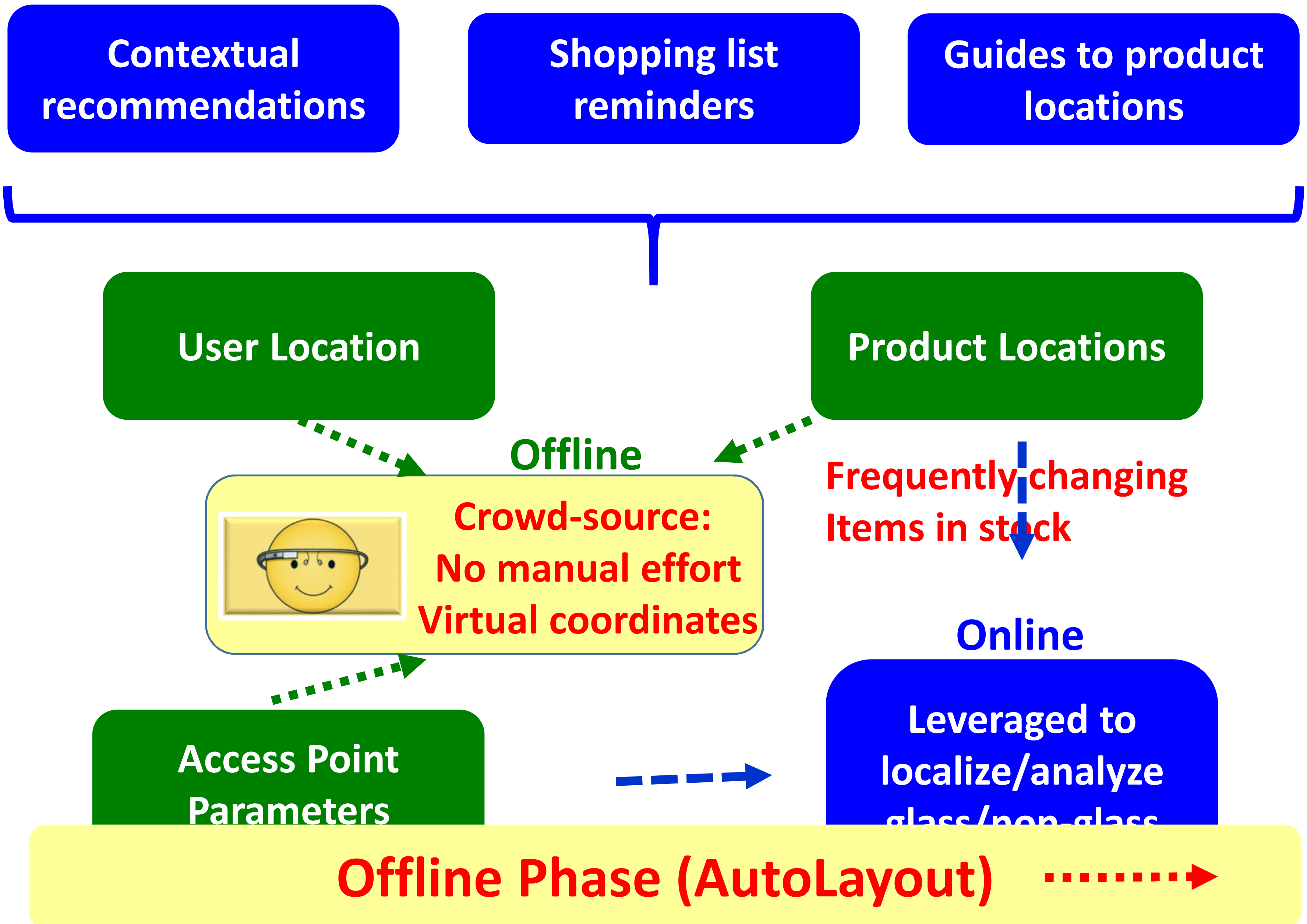
**Build product map**

+

**AP locations**

**Online**

**Localize Non-Glass users**


Contextual recommendations
Shopping list reminders
Guides to product locations
User Location
Product Locations
Offline
Crowd-source:
No manual effort
Virtual coordinates
Frequently changing
Items in stock
Online
Access Point Parameters
Leveraged to localize/analyze
Offline Phase (AutoLayout)

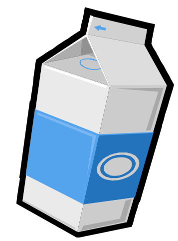
# Problem formulation: unknowns

Localization

$\boldsymbol{sLoc_k^i}$: 2D location of $k^{th}$ shopper after $i^{th}$ step

Product layout

$\boldsymbol{pLoc_j}$: 2D location of $j^{th}$ product in store



$< \boldsymbol{P_l, aLoc_l, \gamma_l} >$: parameters of $l^{th}$ *access point*

Path loss constant

Access point location

Transmit power

Critical to localize non-glass users!

# Problem formulation

**Minimize**

**Leveraged by [EZ]: ties shopper locations with AP parameters (AP location, AP Tx power, ..)**

Wi-Fi term — Inertial sensor term — Camera term

$$\mathbf{w_1} \cdot \boxed{\mathbf{r}(\mathrm{sLoc}, \mathrm{aLoc}, \boldsymbol{P}, \boldsymbol{\gamma})} + \mathbf{w_2} \,.\, \mathrm{p}(\mathrm{sLoc}, \mathrm{t}) + \mathbf{w_3} \cdot \mathbf{q}(\mathbf{sLoc}, \mathbf{pLoc})$$

**Log Distance Path Loss (LDPL) model:** $\mathrm{eRSS}(sLoc, apLoc, apTxPwr, \gamma) = apTxPwr - 10\,\gamma \log(\| sLoc - apLoc \|)$

**Minimizes error in measured RSS values and those estimated by parameters describing the LDPL model**

$$\sum_{l} \sum_{mRSS_{l,k,i} \in W_l} \left\| mRSS_{l,k,i} - eRSS(sLoc_k^i, aLoc_l, P_l, \gamma_l) \right\|$$

$l^{th}$ AP — Over all measured RSS values from that AP across all users

# Incorporate mobility: inertial sensors

**Minimize**

Wi-Fi term | Inertial sensor term | Camera term

$$\mathbf{w_1} \cdot \mathbf{r}(\mathrm{sLoc}, \mathrm{aLoc}, \boldsymbol{P}, \boldsymbol{\gamma}) + \mathbf{w_2} \,.\, \boxed{\mathrm{p}(\mathrm{sLoc}, \mathrm{t})} + \mathbf{w_3} \cdot \mathbf{q}(\mathbf{sLoc}, \mathbf{pLoc})$$

**Accelerometer:** step-count [Zee, UnLoc] → distance

**Compass:** heading direction

**For all shoppers, at all steps:**

$$- \boldsymbol{x_{i+1} \approx x_i + d * cos(\theta)}$$

$$- \boldsymbol{y_{i+1} \approx y_i + d * sin(\theta)}$$

Displacement between inferred locations at consecutive steps close to the estimate of displacement from the inertial sensors

$$\sum_i \sum_k \left\| \boldsymbol{sLoc_k^{i+1} - sLoc_k^i - \hat{e}_k^i} \right\|^2$$

$i^{th}$step $k^{th}$shopper

$$\hat{e}_k^i = \left[\cos\theta_k^i \;\; \sin\theta_k^i\right]^T$$

# Tie in product locations: camera

**Minimize**

Wi-Fi term — Inertial sensor term — Camera term

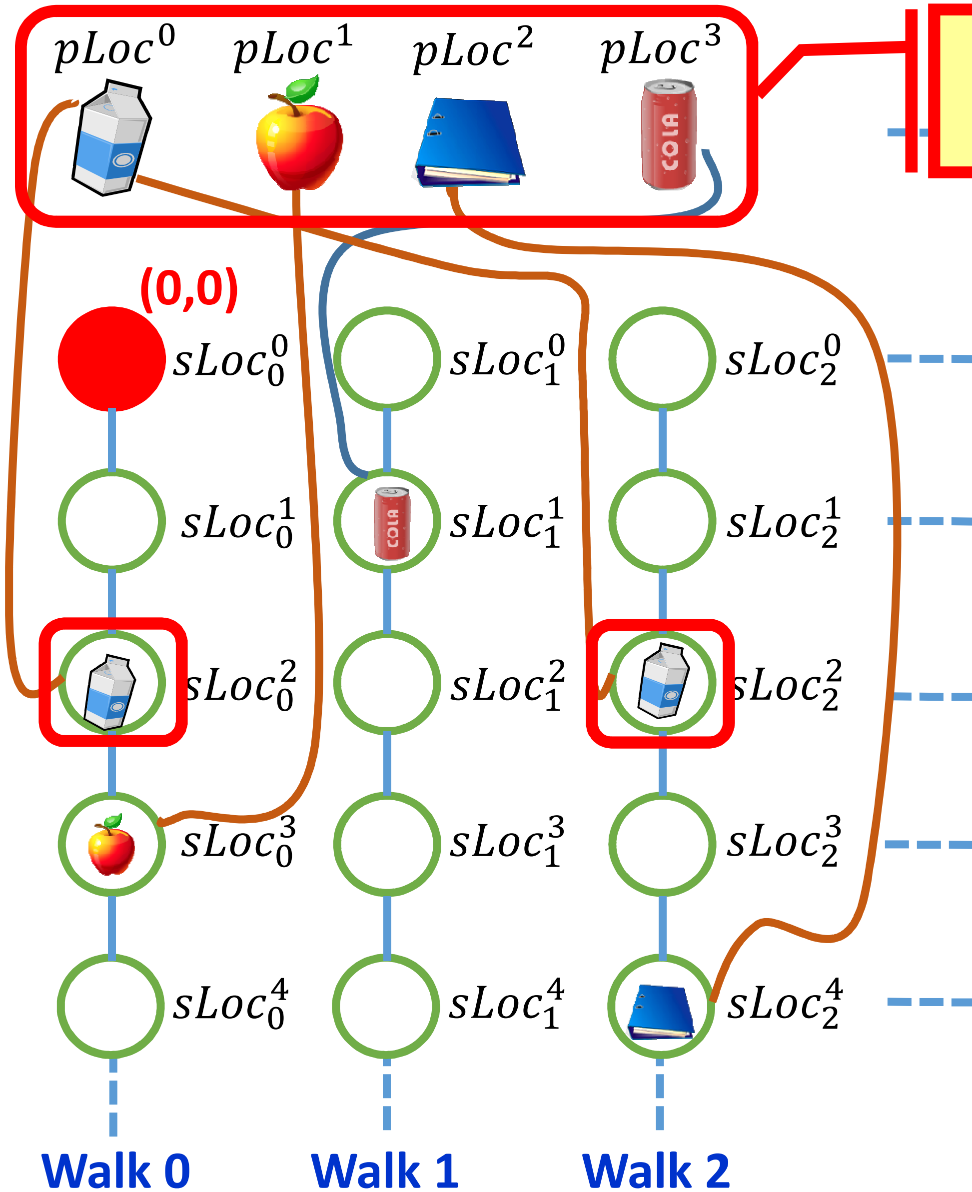
$$\mathbf{w_1}\cdot\mathbf{r}(\mathrm{sLoc},\mathrm{aLoc},\boldsymbol{P},\boldsymbol{\gamma})+\mathbf{w_2}\,.\,\mathrm{p}(\mathrm{sLoc},\mathrm{t})+\mathbf{w_3}\cdot\boxed{\mathbf{q}(\mathbf{sLoc},\mathbf{pLoc})}$$

**Leverage Google Reverse Image search to obtain labels for product images**

**All shopper locations from where a particular product was seen must be close to each other**

$$\underbrace{\sum_j}_{j^{th}\text{product}}\underbrace{\sum_{<k,m>\in L_j}}_{k^{th}\text{shopper saw } j^{th}\text{product at } m^{th}\text{step}}\left\|sLoc_k^m-pLoc_j\right\|^2$$

# Optimization




**Leverage product locations to align different walks**

- Origin: $sLoc_0^0 = (0,0)$
- Leverage mobility: Locations within a walk are connected via inertial sensor data

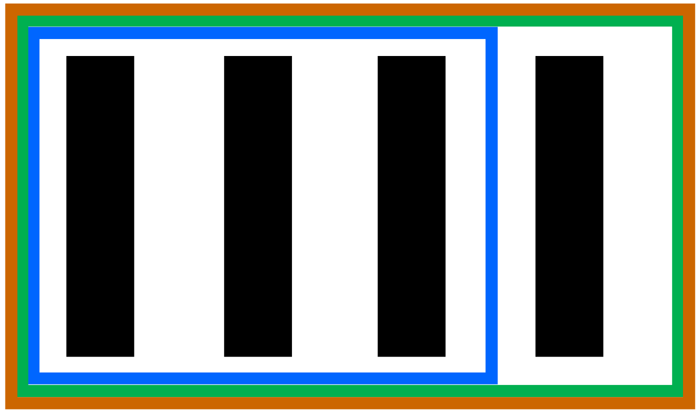
**Milk seen by shopper 0 and shopper 2**

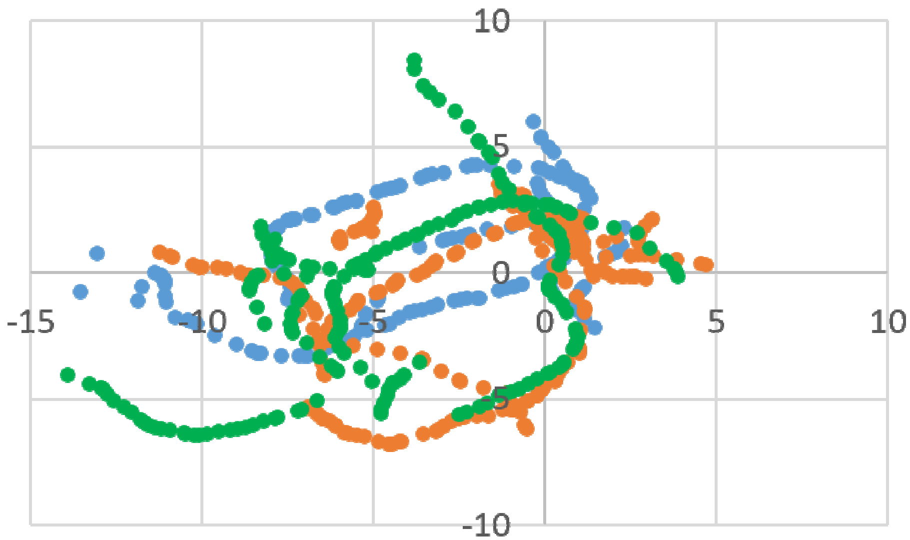
- Run BFS to initialize all shopper & product locations

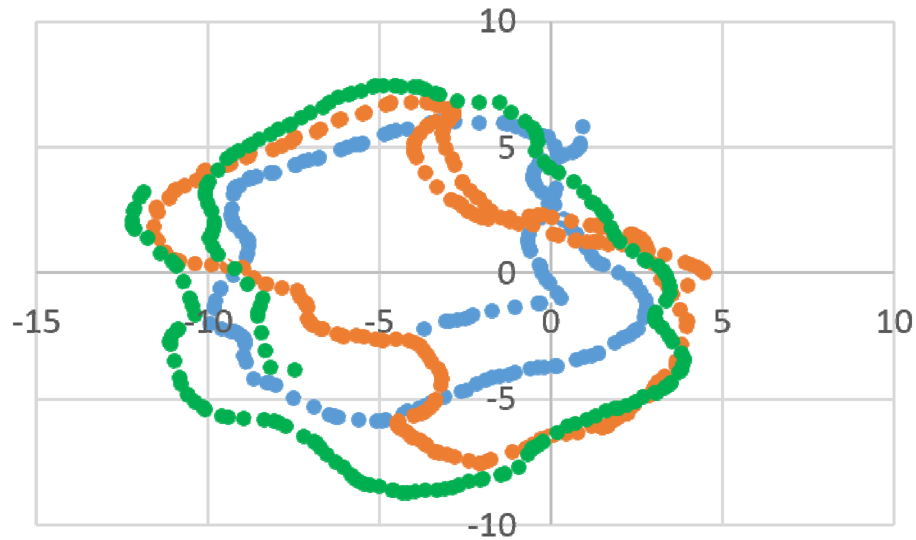
- Gradient descent: refine initial estimates

# Example walks around aisles in Target

Actual walks around aisles



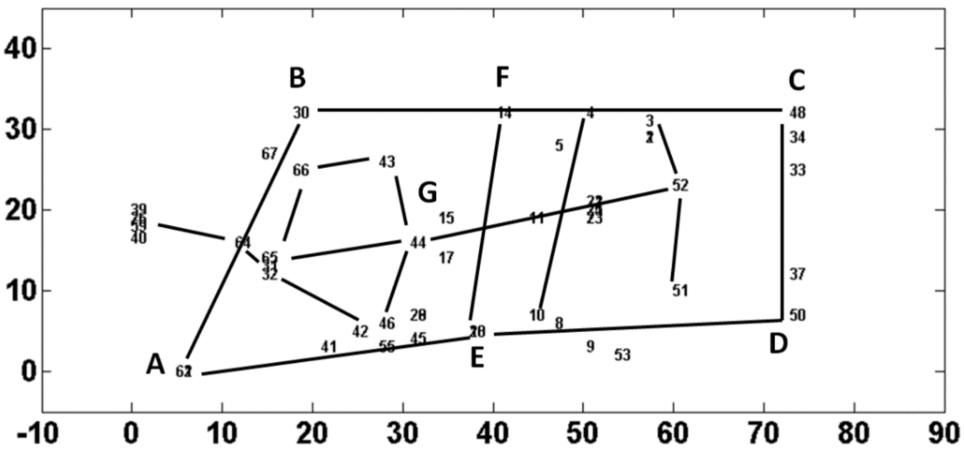


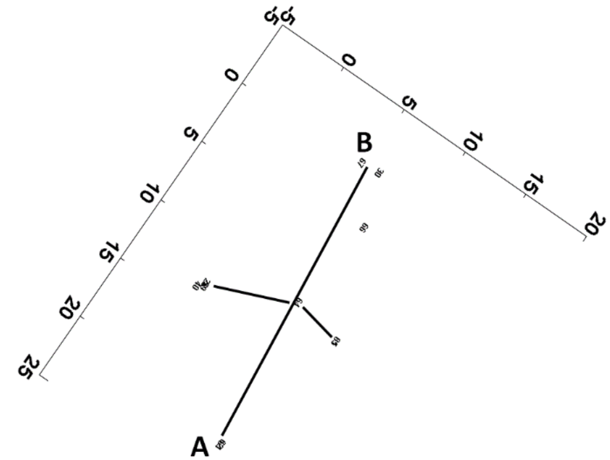

**After BFS: all tracks are in same coordinate system**

**After optimization: tracks look closer to actual walk**

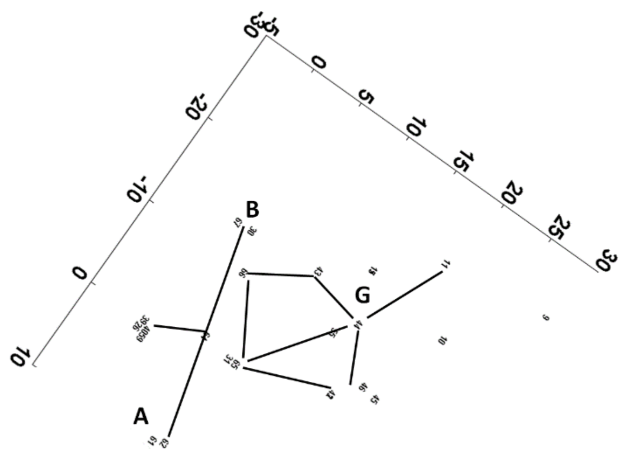
# Inferred layout for H-E-B: improves with more shoppers



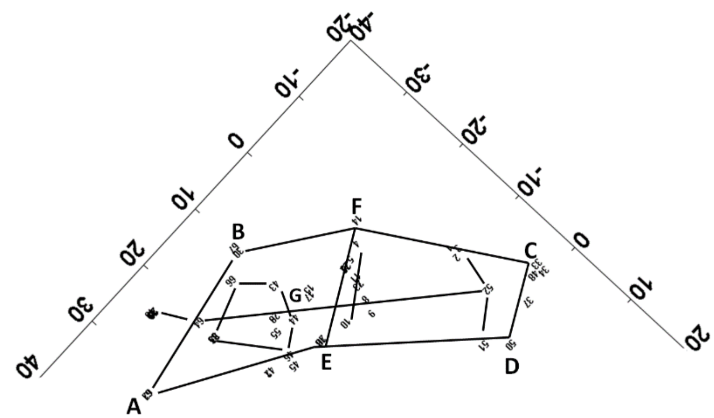

Actual Layout




2 Shoppers




4 shoppers

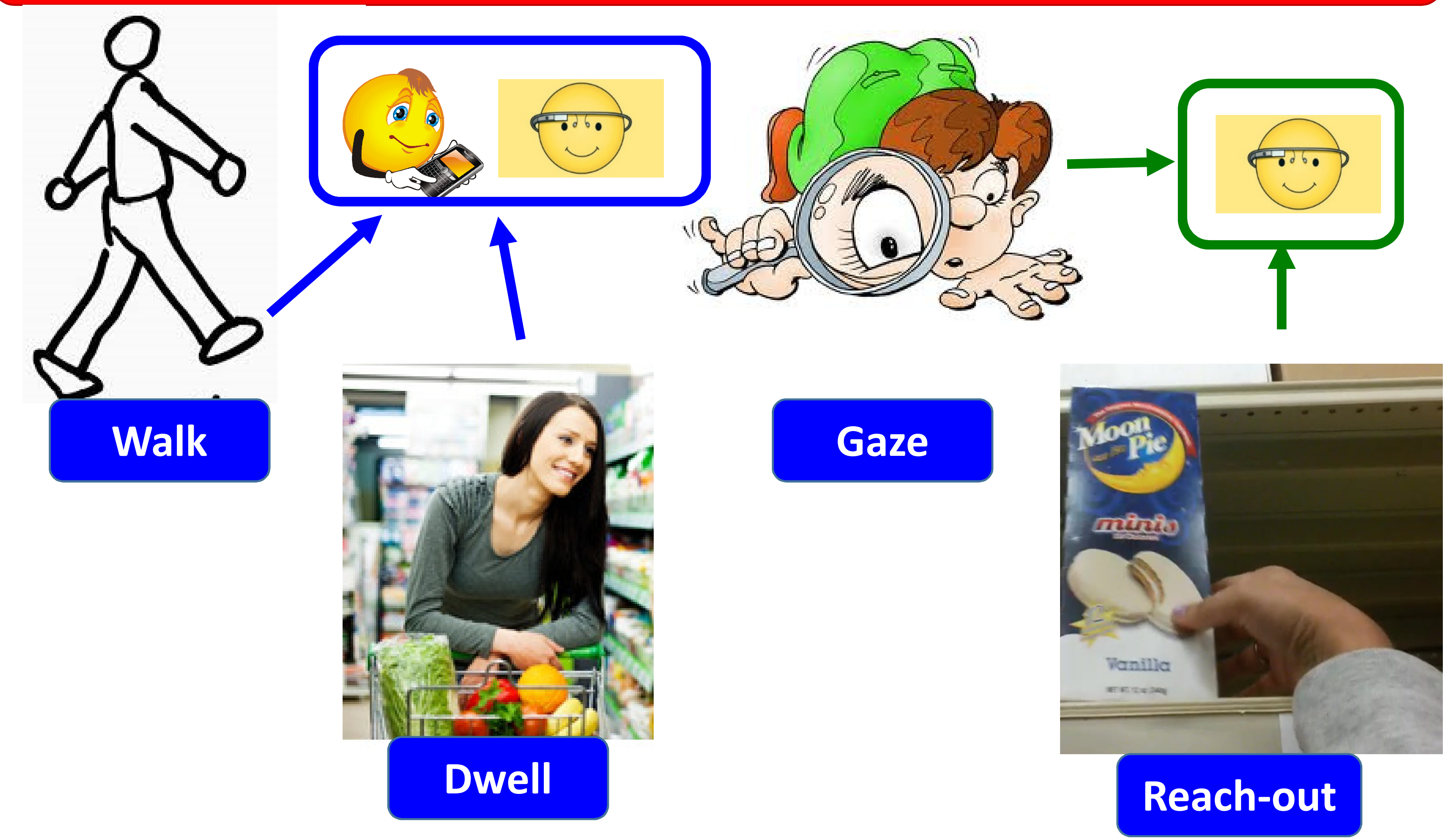



7 shoppers

AutoLayout

Behavior Classification

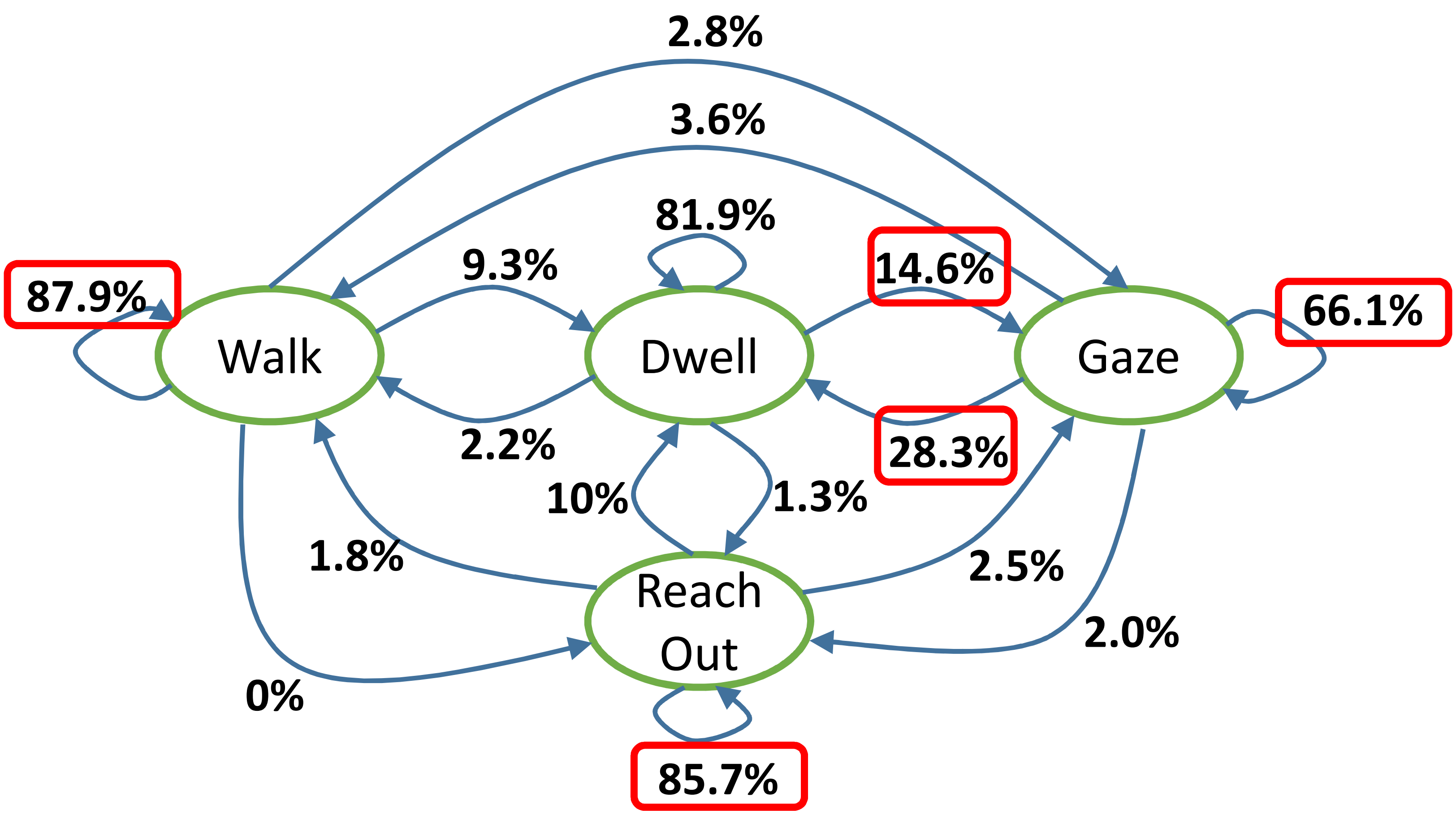
# Overview: ThirdEye - User analytics




**In a retail setting**

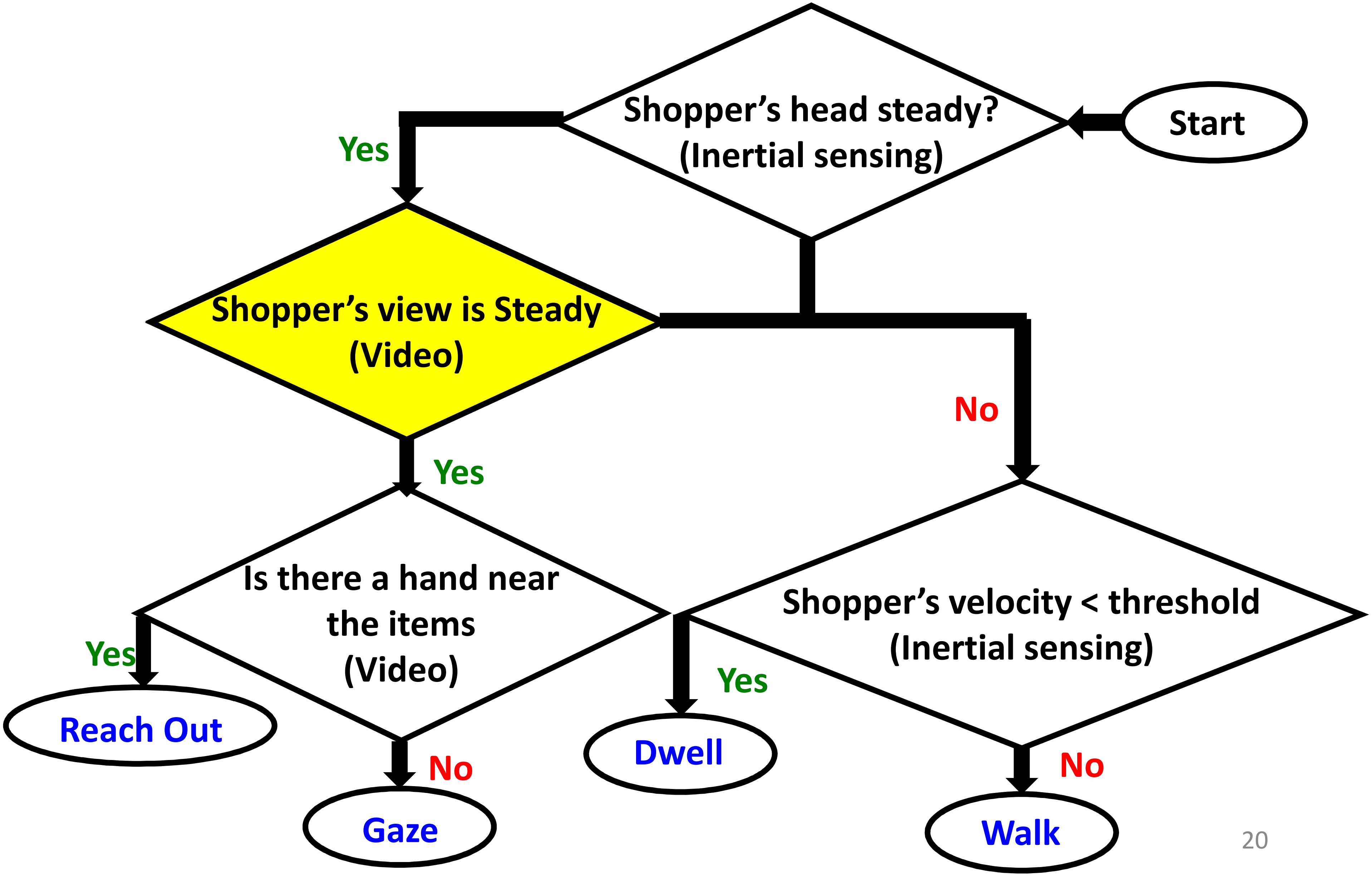
**Analyzed 3 hours of shopping videos from 7 shoppers wearing Google Glass from 2 large stores: H-E-B and Target**

**Time spent: dwell (50.7%) > gaze (23.7%) > walk (17.3%) > reach-out (8.2%)**




- **Most frequent inter-state transitions: dwell and gaze**
- **All states: tend to remain in same state for few seconds**

# Behavior classification algorithm



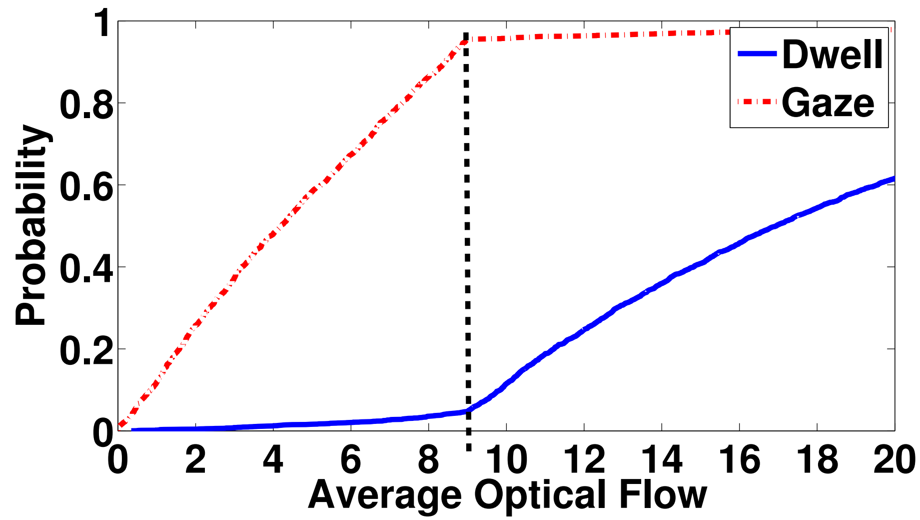
# Gaze and Reach-out

When shopper is gazing/reaching-out scene in front of him does not change

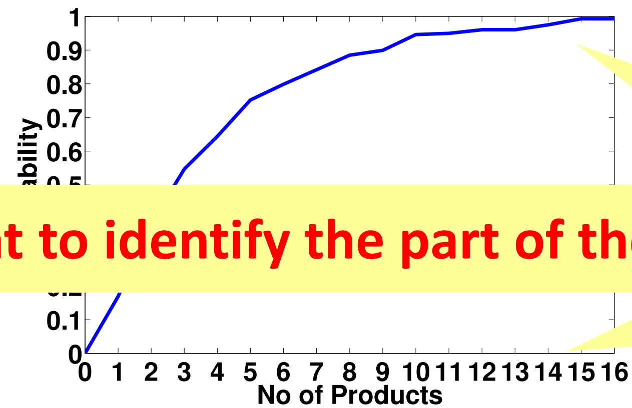
**Leverage vision based technique Optical Flow to detect gaze/reach-out**

- Optical flow (of): difference in terms of pixels between consecutive images
- If $of < of_{gaze}$ detect gaze/reach-out

**88% detection rate at 1.8% false detections**



# Attention identification



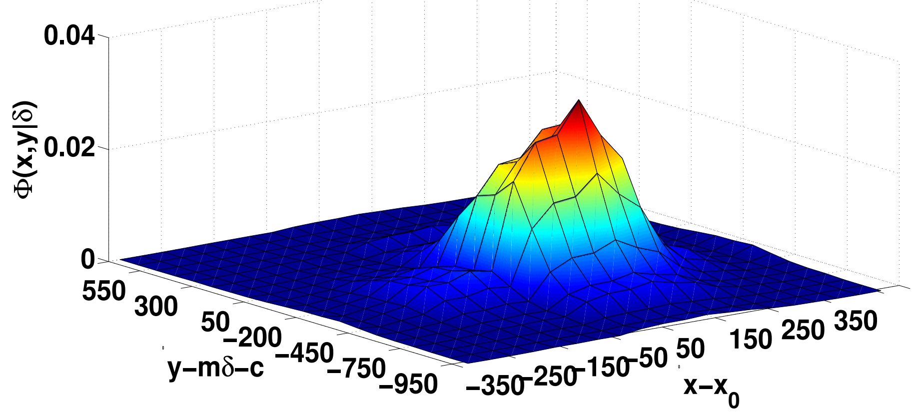

**A frame may**

**Important to identify the part of the frame that user was interested in!**

**as 16 items!**

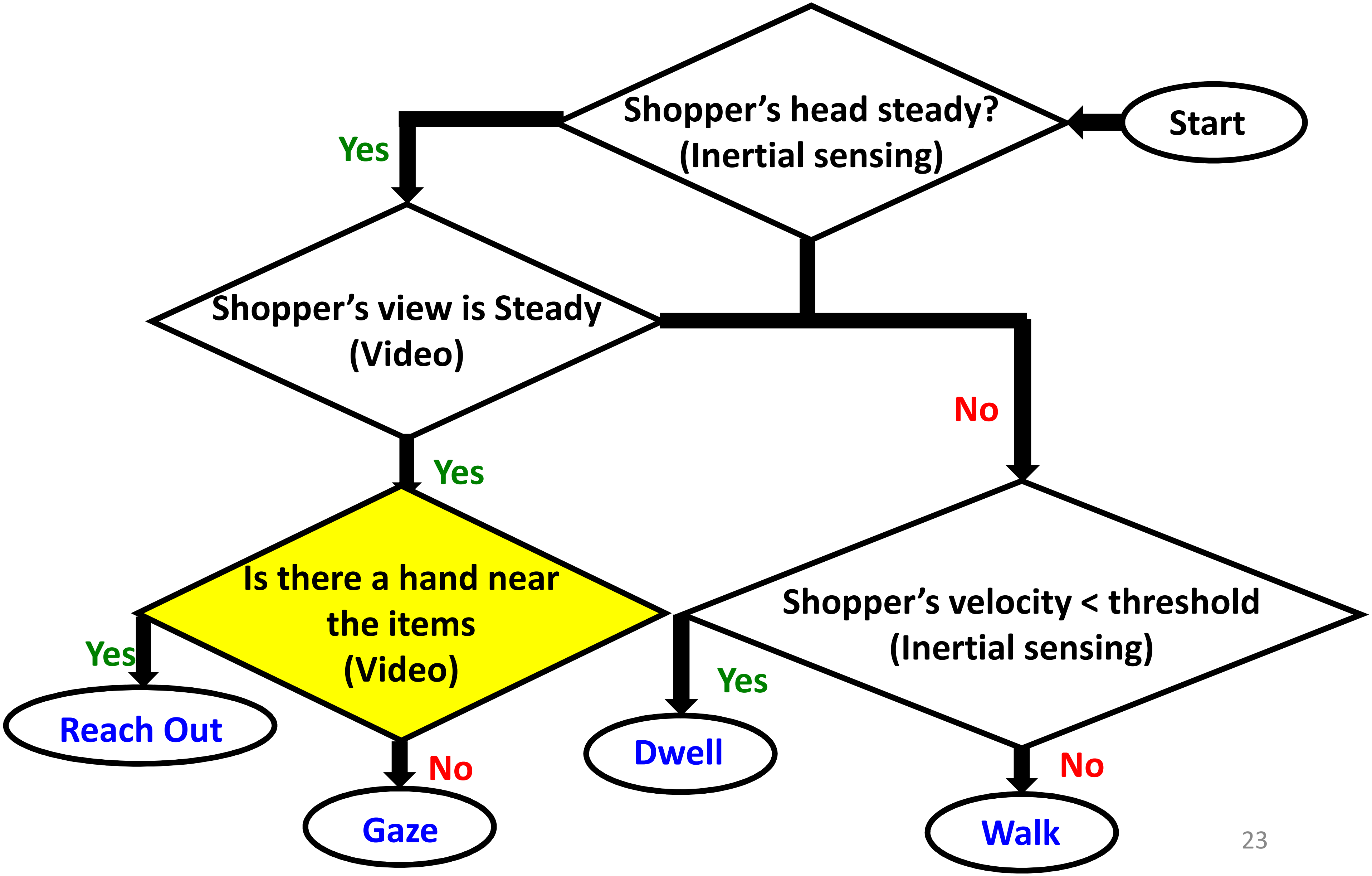
- **Estimated X: center of frame with offset for camera position**
- **Estimated Y: Function of head tilt**

**PDF of error in X, Y estimates**




**Accuracy of product identification: Top 1: 76%, Top 3: 90%**

# Behavior classification algorithm



# Reach-out detection

**Reach-out indicates high degree of interest: important to detect**


**Hand seen in the frame:**

— **detect hands to detect reach-out**

**Train TextonBoost Classifier**

# Leveraging TextonBoost classifier

**Divided hand**

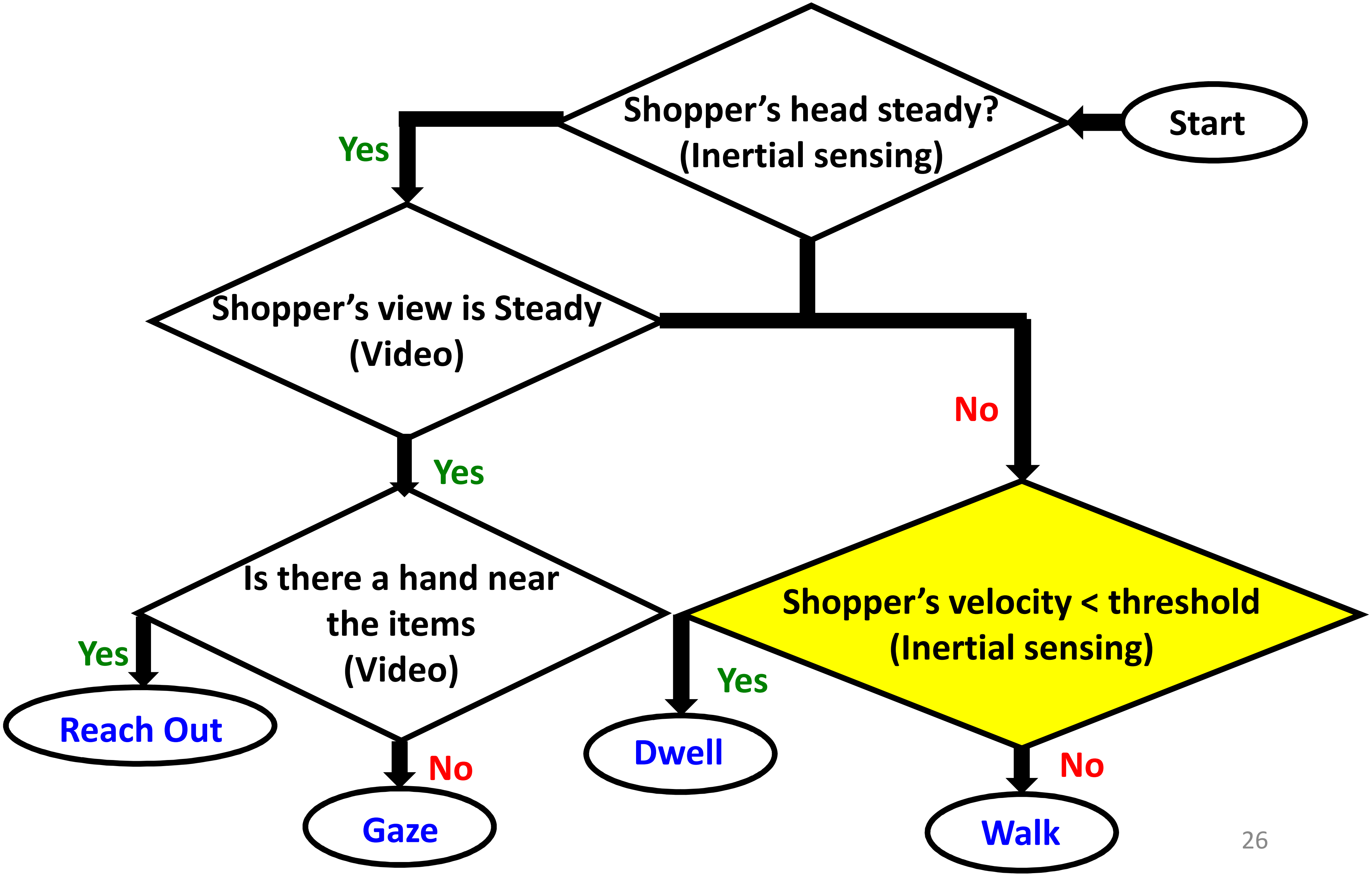

**Cluster together nearby segments**

**Spurious hand**


**Ignore very small segments**

**Detection success rate: 86% False detection rate: 15%**

# Behavior classification algorithm



# Dwell detection

**Accelerometer showing that user is static?**

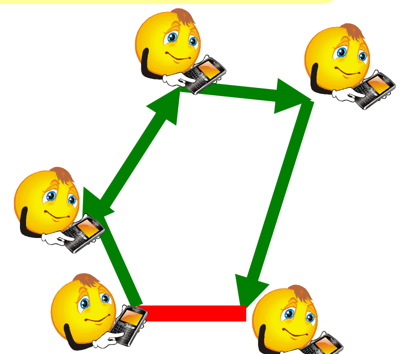
Shopper may not be static, he may take few steps looking at nearby items

**Dwell characterized by small net displacement!**

**Detect dwell based on periods of low net displacement**

Suppose K steps in time window $\tau$ and heading at step *i* is $\theta_i$

— Detect steps using prior work Zee [MobiCom 2011]



**Net displacement in 5 sec**

Magnitude of net velocity vector

$$\|v\| = \sqrt{\left(\sum_{i=1}^{i=K} \cos\theta_i\right)^2 + \left(\sum_{i=1}^{i=K} \sin\theta_i\right)^2}$$

**Dwell if:**

$\|v\| < \|v\|_{dwell}$

**95% detection rate at 10% false alarms**

# Related work

**Research**

- Localization [MobiSys12, MobiCom10, MobiCom12, ..]
- Vision [ICCV09, ICCV13, ..]
- Robotics [UbiComp09, ..]
- Human-activity sensing [Sensys07, ..]
- Shopping-behavior [UbiCom08, Pervasive Computing11, ..]

**Industry**


iBeacon

Get In-Store Notifications


Point Inside

# Conclusion

## Our contributions

— Fuse Wi-Fi, inertial sensor and video data from smart glasses
— AutoLayout: Map the store without any user or store input
— Use these inferences to track glass/non-glass users in online phase
— Characterize walk, dwell, gaze and reaching-out activities of shoppers
— Attention identification within the captured frame

## Future work

— Larger data-set for patterns representative of more diverse population
— In-depth analytics of shoppers

# Thank You!

swati@cs.utexas.edu